**Problem Statement:**

In the rapidly evolving financial landscape, predicting stock prices is crucial for informed investment decisions. This research aims to develop robust predictive models for Tata Consultancy Services (TCS) by leveraging historical stock price data and analyzing sentiment from relevant news articles. By integrating quantitative and qualitative data, we seek to enhance the accuracy of stock price movement predictions, providing valuable insights for investors.

**XGBoost**

XGBoost (eXtreme Gradient Boosting) is a popular open-source gradient boosting framework that is widely used for classification and regression tasks. Its importance lies in its ability to handle large datasets and provide high accuracy results.

**Key Parameters:**

* **n\_estimators**: The number of trees in the ensemble.
* **max\_depth**: The maximum depth of each tree.
* **learning\_rate**: The step size of each iteration while moving toward a minimum of a loss function.
* **gamma**: The minimum loss reduction required to make a further partition on a leaf node of the tree.
* **subsample**: The fraction of samples to be used for fitting the individual base learners.
* **colsample\_bytree**: The fraction of columns to be randomly sampled for each tree.

**Light GBM**

Light GBM is another popular gradient boosting framework that is designed to be highly efficient and scalable. Its importance lies in its ability to handle large datasets and provide high accuracy results.

**Key Parameters:**

* **num\_leaves**: The number of leaves in the decision tree.
* **max\_depth**: The maximum depth of the decision tree.
* **learning\_rate**: The step size of each iteration while moving toward a minimum of a loss function.
* **n\_estimators**: The number of trees in the ensemble.
* **min\_data\_in\_leaf**: The minimum number of data points required to be in a leaf.
* **min\_sum\_hessian\_in\_leaf**: The minimum sum of the Hessian matrix required to be in a leaf.

**ChatGPT 3.5**

ChatGPT 3.5 is a conversational AI model developed by OpenAI that is capable of generating human-like text responses. Its importance lies in its ability to understand and respond to natural language inputs.

**Key Parameters:**

* **temperature**: The temperature parameter controls the randomness of the generated text.
* **top\_p**: The top\_p parameter controls the proportion of the most likely tokens to be considered.
* **max\_length**: The maximum length of the generated text.

**Facebook BART-large-cnn**

BART-large-cnn is a pre-trained language model developed by Facebook that is capable of generating text summaries. Its importance lies in its ability to understand and summarize long pieces of text.

**Key Parameters:**

* **max\_length**: The maximum length of the generated summary.
* **min\_length**: The minimum length of the generated summary.
* **num\_beams**: The number of beams to use for beam search.

**NLTK Library**

NLTK (Natural Language Toolkit) is a popular Python library used for natural language processing tasks. Its importance lies in its ability to provide a wide range of tools and resources for text processing.

**Key Parameters:**

* **tokenize**: The tokenize function is used to split text into individual words or tokens.
* **stem**: The stem function is used to reduce words to their base form.
* **lemmatize**: The lemmatize function is used to reduce words to their base form using a dictionary.

Here is a sample c